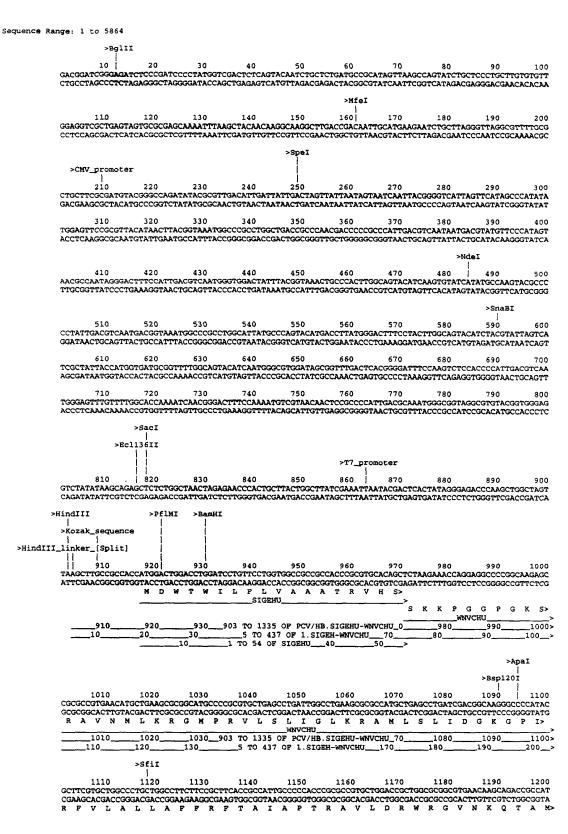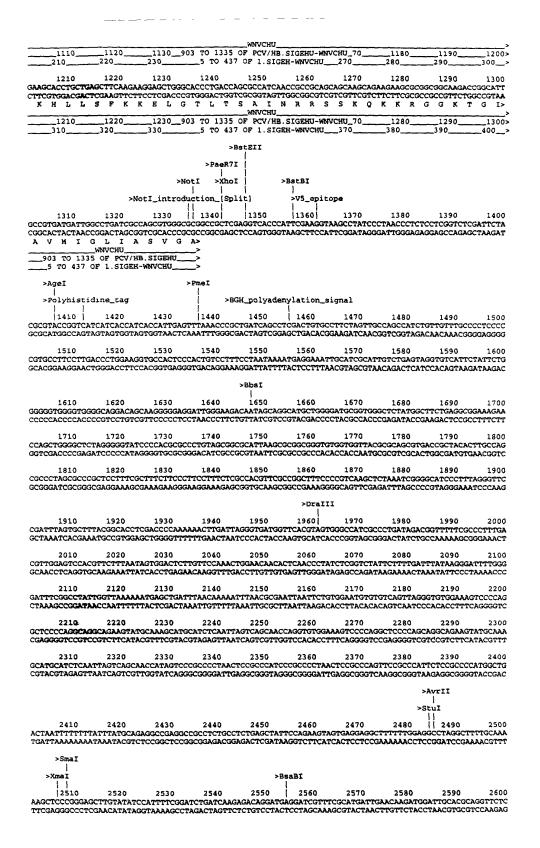Docket Number: UPN-4105 (M2351)
Title: Compositions and Methods of Using Capsid
Protein from Flaviviruses and Pestiviruses
Inventors: David B. Weiner and Joo-Sung Yang
Sheet 4 of 30

Sequence Range: 1 to 5864

```
                   >BglII
                     |
       10  |    20        30        40        50        60        70        80        90       100
GACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT
CTGCCTAGCCCTCTAGAGGGCTAGGGGATACCAGCTGAGAGTCATGTTAGACGAGACTACGGCGTATCAATTCGGTCATAGACGAGGGACGAACACACAA


                                                                      >MfeI
                                                                        |
       110       120       130       140       150       160|      170       180       190       200
GGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCG
CCTCCAGCGACTCATCACGCGCTCGTTTTAAATTCGATGTTGTTCCGTTCCGAACTGGCTGTTAACGTACTTCTTAGACGAATCCCAATCCGCAAAACGC


                                               >SpeI
                                                 |
          >CMV_promoter                          |
          |
       210       220       230       240       250       260       270       280       290       300
CTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
GACGAAGCGCTACATGCCCGGTCTATATGCGCAACTGTAACTAATAACTGATCAATAATTATCATTAGTTAATGCCCCAGTAATCAAGTATCGGGTATAT


       310       320       330       340       350       360       370       380       390       400
TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
ACCTCAAGGCGCAATGTATTGAATGCCATTTACCGGGCGGACCGACTGGCGGGTTGCTGGGGGCGGGTAACTGCAGTTATTACTGCATACAAGGGTATCA


                                                                              >NdeI
                                                                                |
       410       420       430       440       450       460       470       480 |     490       500
AACGCCAATAGGGACTTTCCATTGACGTCAATGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC
TTGCGGTTATCCCTGAAAGGTAACTGCAGTTACCCACCTGATAAATGCCATTTGACGGGTGAACCGTCATGTAGTTCACATAGTATACGGTTCATGCGGG


                                                                              >SnaBI
                                                                                |
       510       520       530       540       550       560       570       580 |     590       600
CCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCA
GGATAACTGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCATGTACTGGAATACCCTGAAAGGATGAACCGTCATGTAGATGCATAATCAGT


       610       620       630       640       650       660       670       680       690       700
TCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAA
AGCGATAATGGTACCACTACGCCAAAACCGTCATGTAGTTACCCGCACCTATCGCCAAACTGAGTGCCCCTAAAGGTTCAGAGGTGGGGTAACTGCAGTT


       710       720       730       740       750       760       770       780       790       800
TGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG
ACCCTCAAACAAAACCGTGGTTTTAGTTGCCCTGAAAGGTTTTACAGCATTGTTGAGGCGGGGTAACTGCGTTTACCCGCCATCCGCACATGCCACCCTC


          >SacI
           |
       >Ecl136II
        | |
        | |
        | |
       810  .| 820       830       840       850       860 |     870       880       890       900
                                                     >T7_promoter
GTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGT
CAGATATATTCGTCTCGAGAGACCGATTGATCTCTTGGGTGACGAATGACCGAATAGCTTTAATTATGCTGAGTGATATCCCTCTGGGTTCGACCGATCA


>HindIII            >PflMI    >BamHI
 |                   |         |
>Kozak_sequence      |         |
 | |                 |         |
>HindIII_linker_[Split]        |
 || |                |         |
       910       920 |     930 |     940       950       960       970       980       990      1000
TAAGCTTGCCGCCACCATGGACTGGACCTGGATCCTGTTCCTGGTGGCCGCCGCCACCCGCGTGCACAGCTCTAAGAAACCAGGAGGCCCCGGCAAGAGC
ATTCGAACGGCGGTGGTACCTGACCTGGACCTAGGACAAGGACCACCGGCGGCGGTGGGCGCACGTGTCGAGATTCTTTGGTCCTCCGGGGCCGTTCTCG
             M  D  W  T  W  I  L  F  L  V  A  A  A  T  R  V  H  S>
                              SIGEHU_____>
                                                          S  K  K  P  G  G  P  G  K  S>
                                                             _____WNVCHU_____>
         ____910_____920_____930___903 TO 1335 OF PCV/HB.SIGEHU-WNVCHU_0_____980_____990_____1000>
         ____10_____20_____30_____5 TO 437 OF 1.SIGEH-WNVCHU____70_____80_____90_____100__>
                            _____10_____1 TO 54 OF SIGEHU___40_____50____>


                                                              >ApaI
                                                                |
                                                           >Bsp120I
                                                             |  |
      1010      1020      1030      1040      1050      1060      1070      1080      1090 |    1100
CGCGCCGTGAACATGCTGAAGCGCGGCATGCCCCGCGTGCTGAGCCTGATTGGCCTGAAGCGCGCCATGCTGAGCCTGATCGACGGCAAGGGCCCCATAC
GCGCGGCACTTGTACGACTTCGCGCCGTACGGGGCGCACGACTCGGACTAACCGGACTTCGCGCGGTACGACTCGGACTAGCTGCCGTTCCCGGGGTATG
 R  A  V  N  M  L  K  R  G  M  P  R  V  L  S  L  I  G  L  K  R  A  M  L  S  L  I  D  G  K  G  P  I>
                                        WNVCHU_____>
         ____1010_____1020_____1030__903 TO 1335 OF PCV/HB.SIGEHU-WNVCHU_70_____1080_____1090_____1100>
         ____110_____120_____130_____5 TO 437 OF 1.SIGEH-WNVCHU____170_____180_____190_____200__>


      >SfiI
        |
      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
GCTTCGTGCTGGCCCTGCTGGCCTTCTTCCGCTTCACCGCCATTGCCCCCACCCGCGCCGTGCTGGACCGCTGGCGCGGCGTGAACAAGCAGACCGCCAT
CGAAGCACGACCGGGACGACCGGAAGAAGGCGAAGTGGCGGTAACGGGGGTGGGCGCGGCACGACCTGGCGACCGCGCCGCACTTGTTCGTCTGGCGGTA
 R  F  V  L  A  L  L  A  F  F  R  F  T  A  I  A  P  T  R  A  V  L  D  R  W  R  G  V  N  K  Q  T  A  M>
```

FIG 4

Docket Number: UPN-4105 (M2351)
Title: Compositions and Methods of Using Capsid
Protein from Flaviviruses and Pestiviruses
Inventors: David B. Weiner and Joo-Sung Yang
Sheet 5 of 30

```
                                               WNVCHU                                                   >
_____1110_____1120_____1130__903 TO 1335 OF PCV/HB.SIGEHU-WNVCHU_70_____1180_____1190_____1200>
_____210_____220_____230_____5 TO 437 OF 1.SIGEH-WNVCHU___270_____280_____290_____300__>

      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300
GAAGCACCTGCTGAGCTTCAAGAAGGAGCTGGGCACCCTGACCAGCGCCATCAACCGCCGCAGCAGCAAGCAGAAGAAGCGCGGCCGGCAAGACCGGCATT
CTTCGTGGACGACTCGAAGTTCTTCCTCGACCCGTGGGACTGGTCGCGGTAGTTGGCGGCGTCGTCGTTCGTCTTCTTCGCGCCGCCGTTCTGGCCGTAA
   K  H  L  L  S  F  K  K  E  L  G  T  L  T  S  A  I  N  R  R  S  S  K  Q  K  K  R  G  G  K  T  G  I>
                                               WNVCHU                                                   >
_____1210_____1220_____1230__903 TO 1335 OF PCV/HB.SIGEHU-WNVCHU_70_____1280_____1290_____1300>
_____310_____320_____330_____5 TO 437 OF 1.SIGEH-WNVCHU___370_____380_____390_____400__>
```

```
                                    >BstEII
                                      |
                           >PaeR7I   |
                             |  |
                    >NotI  >XhoI  |         >BstBI
                     |      |   |           |
         >NotI_introduction_[Split]      >V5_epitope
                    | |          |           |
    1310      1320    1330  | | 1340|   |1350   |1360|   1370      1380      1390      1400
GCCGTGATGATTGGCCTGATCGCCAGCGTGGGCGCGGCCGCTCGAGGTCACCCATTCGAAGGTAAGCCTATCCCTAACCCTCCTCGGTCTCGATTCTA
CGGCACTACTAACCGGACTAGCGGTCGCACCCGCGCCGGCGAGCTCCAGTGGGTAAGCTTCCATTCGGATAGGGATTGGGAGAGGAGCCAGAGCTAAGAT
  A  V  M  I  G  L  I  A  S  V  G  A>
                  WNVCHU                >
____903 TO 1335 OF PCV/HB.SIGEHU____>
____5 TO 437 OF 1.SIGEH-WNVCHU_____>
```

```
    >AgeI                   >PmeI
     |                       |
    >Polyhistidine_tag       |         >BGH_polyadenylation_signal
     | |                     |          |
     | |1410 |   1420      1430  |1440    1450   | 1460    1470      1480      1490      1500
CGCGTACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCC
GCGCATGGCCAGTAGTAGTGGTAGTGGTAACTCAAATTTGGGCGACTAGTCGGAGCTGACACGGAAGATCAACGGTCGGTAGACAACAAACGGGGAGGGG

      1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
CGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCCATTGTCTGAGTAGGTGTCATTCTATTCTG
GCACGGAAGGAACTGGGACCTTCCACGGTGAGGGTGACAGGAAAGGATTATTTTACTCCTTTAACGTAGCGTAACAGACTCATCCACAGTAAGATAAGAC
```

```
                                    >BbsI
                                      |
      1610      1620      1630      1640      1650      1660      1670      1680      1690      1700
GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAA
CCCCCACCCCACCCCGTCCTGTCGTTCCCCCTCCTAACCCTTCTGTTATCGTCCGTACGACCCCTACGCCACCCGAGATACCGAAGACTCCGCCTTTCTT

      1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
CCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGTACACTTGCCAG
GGTCGACCCCGAGATCCCCCATAGGGGTGCGCGGGACATCGCCGCGTAATTCGCGCCGCCCACACCACCAATGCGCGTCGCACTGGCGATGTGAACGGTC

      1810      1820      1830      1840      1850      1860      1870      1880      1890      1900
CGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGGGTTC
GCGGGATCGCGGGCGAGGAAAGCGAAAGAAGGGAAGGAAAGAGCGGTGCAAGCGGCCGAAAGGGGCAGTTCGAGATTTAGCCCCGTAGGGAAATCCCAAG
```

```
                                    >DraIII
                                      |
      1910      1920      1930      1940      1950      1960|     1970      1980      1990      2000
CGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGA
GCTAAATCACGAAATGCCGTGGAGCTGGGGTTTTTTGAACTAATCCCACTACCAAGTGCATCACCCGGTAGCGGGACTATCTGCCAAAAAGCGGGAAACT

      2010      2020      2030      2040      2050      2060      2070      2080      2090      2100
CGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGGG
GCAACCTCAGGTGCAAGAAATTATCACCTGAGAACAAGGTTTGACCTTGTTGTGAGTTGGGATAGAGCCAGATAAGAAAACTAAATATTCCCTAAAACCC

      2110      2120      2130      2140      2150      2160      2170      2180      2190      2200
GATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAG
CTAAAGCCGGATAACCAATTTTTTTACTCGACTAAATTGTTTTTAAATTGCGCTTAATTAAGACACCTTACACACAGTCAATCCCACACCTTTCAGGGGTC

      2210      2220      2230      2240      2250      2260      2270      2280      2290      2300
GCTCCCCAGGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAA
CGAGGGGGTCCGTCCGTCTTCATACGTTTCGTACGTAGAGTTAATCAGTCGTTGGTCCACACCTTTCAGGGGTCCGAGGGGTCGTCCGTCTTCATACGTTT

      2310      2320      2330      2340      2350      2360      2370      2380      2390      2400
GCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCCATGGCTG
CGTACGTAGAGTTAATCAGTCGTTGGTATCAGGGCGGGGATTGAGGCGGGTAGGGCGGGGATTGAGGCGGGTCAAGGCGGGTAAGAGGCGGGGTACCGAC
```

```
                                                                      >AvrII
                                                                       |
                                                                     >StuI
                                                                      | |
      2410      2420      2430      2440      2450      2460      2470  | 2480 || 2490      2500
ACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTGCAAA
TGATTAAAAAAAAТAAATACGTCTCCGGCTCCGGCGGAGACGGAGACTCGATAAGGTCTTCATCACTCCTCCGAAAAAACCTCCGGATCCGAAAACGTTT
```

```
    >SmaI
     |
    >XmaI                                       >BsaBI
     | |                                          |
     | |2510      2520      2530      2540      2550  | 2560      2570      2580      2590      2600
AAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTC
TTCGAGGGCCCTCGAACATATAGGTAAAAGCCTAGACTAGTTCTCTGTCCTACTCCTAGCAAAGCGTACTAACTTGTTCTACCTAACGTGCGTCCAAGAG
```

Docket Number: UPN-4105 (M2351)
Title: Compositions and Methods of Using Capsid
Protein from Flaviviruses and Pestiviruses
Inventors: David B. Weiner and Joo-Sung Yang
Sheet 6 of 30

```
                                                                              >EheI
                                                                              |
                                                                            >NarI
                                                                            ||
                                                                          >KasI
                                                                          |||
        2610      2620      2630      2640      2650      2660      2670      2680      2690      |||2700
CGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCC
GCCGGCGAACCCACCTCTCCGATAAGCCGATACTGACCCGTGTTGTCTGTTAGCCGACGAGACTACGGCGGCACAAGGCCGACAGTCGCGTCCCCGCGGG

                                            >PstI                           >MscI
                                            |                               |
        2710      2720      2730      2740      2750      2760      2770      2780      2790      2800
GGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCA
CCAAGAAAAACAGTTCTGGCTGGACAGGCCACGGGACTTACTTGACGTCCTGCTCCGTCGCGCCGATAGCACCGACCGGTGCTGCCCGCAAGGAACGCGT

            >Tth111I
            |
        2810      2820      2830      2840      2850      2860      2870      2880      2890      2900
GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGA
CGACACGAGCTGCAACAGTGACTTCGCCCTTCCCTGACCGACGATAACCCGCTTCACGGCCCCGTCCTAGAGGACAGTAGAGTGGAACGAGGACGGCTCT

        2910      2920      2930      2940      2950      2960      2970      2980      2990      3000
AAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACG
TTCATAGGTAGTACCGACTACGTTACGCCGCCGACGTATGCGAACTAGGCCGATGGACGGGTAAGCTGGTGGTTCGCTTTGTAGCGTAGCTCGCTCGTGC

        3010      3020      3030      3040      3050      3060      3070      3080      3090      3100
TACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATG
ATGAGCCTACCTTCGGCCAGAACAGCTAGTCCTACTAGACCTGCTTCTCGTAGTCCCCGAGCGCGGTCGGCTTGACAAGCGGTCCGAGTTCCGCGCGTAC

        3110      3120      3130      3140      3150      3160      3170      3180      3190      3200
CCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGC
GGGCTGCCGCTCCTAGAGCAGCACTGGGTACCGCTACGGACGAACGGCTTATAGTACCACCTTTTACCGGCGAAAAGACCTAAGTAGCTGACACCGGCCG

            >RsrII
            |
        3210|     3220      3230      3240      3250      3260      3270      3280      3290      3300
TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGG
ACCCACACCGCCTGGCGATAGTCCTGTATCGCAACCGATGGGCACTATAACGACTTCTCGAACCGCCGCTTACCCGACTGGCGAAGGAGCACGAAATGCC

        3310      3320      3330      3340      3350      3360      3370      3380      3390      3400
TATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGCGAAATGACCGACCAAGCGAC
ATAGCGGCGAGGGCTAAGCGTCGCGTAGCGGAAGATAGCGGAAGAACTGCTCAAGAAGACTCGCCCTGAGACCCCAAGCGCTTTACTGGCTGGTTCGCTG

        3410      3420      3430      3440      3450      3460      3470      3480      3490      3500
GCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAG
CGGGTTGGACGGTAGTGCTCTAAAGCTAAGGTGGCGGCGGAAGATACTTTCCAACCCGAAGCCTTAGCAAAAGGCCCTGCGGCCGACCTACTAGGAGGTC

        3510      3520      3530      3540      3550      3560      3570      3580      3590      3600
CGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATA
GCGCCCCTAGAGTACGACCTCAAGAAGCGGGTGGGGTTGAACAAATAACGTCGAATATTACCAATGTTTATTTCGTTATCGTAGTGTTTAAAGTGTTTAT

            >BsmI                                   >Bst1107I
            |                                       |
        3610      3620      3630      3640      3650      3660      3670      3680      3690      3700
AAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTA
TTCGTAAAAAAAGTGACGTAAGATCAACACCAAACAGGTTTGAGTAGTTACATAGAATAGTACAGACATATGGCAGCTGGAGATCGATCTCGAACCGCAT

        3710      3720      3730      3740      3750      3760      3770      3780      3790      3800
ATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAA
TAGTACCAGTATCGACAAAGGACACACTTTAACAATAGGCGAGTGTTAAGGTGTGTTGTATGCTCGGCCTTCGTATTTCACATTTCGGACCCCACGGATT

        3810      3820      3830      3840      3850      3860      3870      3880      3890      3900
TGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGG
ACTCACTCGATTGAGTGTAATTAACGCAACGCGAGTGACGGGCGAAAGGTCAGCCCTTTGGACAGCACGGTCGACGTAATTACTTAGCCGGTTGCGCGCC

        3910      3920      3930      3940      3950      3960      3970      3980      3990      4000
GGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAA
CCTCTCCGCCAAACGCATAACCCGCGAGAAGGCGAAGGAGCGAGTGACTGAGCGACGCGAGCCAGCAAGCCGACGCCGCTCGCCATAGTCGAGTGAGTTT

        4010      4020      4030      4040      4050      4060      4070      4080      4090      4100
GGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTG
CCGCCATTATGCCAATAGGTGTCTTAGTCCCCTATTGCGTCCTTTCTTGTACACTCGTTTTCCGGTCGTTTTCCGGTCCTTGGCATTTTTCCGGCGCAAC

        4110      4120      4130      4140      4150      4160      4170      4180      4190      4200
CTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG
GACCGCAAAAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACCGCTTTGGGCTGTCCTGATATTTCTATGGTCC

        4210      4220      4230      4240      4250      4260      4270      4280      4290      4300
CGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC
GCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGGACGGCGAATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAG

        4310      4320      4330      4340      4350      4360      4370      4380      4390      4400
TCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTA
AGTTACGAGTGCGACATCCATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCGGAAT

        4410      4420      4430      4440      4450      4460      4470      4480      4490      4500
TCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
AGGCCATTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTCCTAATCGTCTCGCTCCATACATCCGC
```